Uncertainty quantification for radiation measurements: Bottom-up error variance estimation using calibration information

T. Burr [a,*], S. Croft [b], T. Krieger [a], K. Martin [a], C. Norman [a], S. Walsh [c]

[a] Nuclear Fuel Cycle Analysis, International Atomic Energy Agency, Box 100, 1400 Vienna, Austria
[b] Nuclear Security and Isotope Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[c] Office of Safeguards Analytical Services, International Atomic Energy Agency, Box 100, 1400 Vienna, Austria

HIGHLIGHTS

- Simulations confirmed a tendency to favour inverse regression for calibration.
- Inverse regression has lower error variance than classical regression followed by inversion.
- Our study extended previous studies in include the case with non-negligible errors in predictors.
- Analytical approximations used to estimate variances are not sufficiently accurate for our application.

ARTICLE INFO

ABSTRACT

One example of top-down uncertainty quantification (UQ) involves comparing two or more measurements on each of multiple items. One example of bottom-up UQ expresses a measurement result as a function of one or more input variables that have associated errors, such as a measured count rate, which individually (or collectively) can be evaluated for impact on the uncertainty in the resulting measured value. In practice, it is often found that top-down UQ exhibits larger error variances than bottom-up UQ, because some error sources are present in the fielded assay methods used in top-down UQ that are not present (or not recognized) in the assay studies used in bottom-up UQ. One would like better consistency between the two approaches in order to claim understanding of the measurement process.

The purpose of this paper is to refine bottom-up uncertainty estimation by using calibration information so that if there are no unknown error sources, the refined bottom-up uncertainty estimate will agree with the top-down uncertainty estimate to within a specified tolerance. Then, in practice, if the top-down uncertainty estimate is larger than the refined bottom-up uncertainty estimate by more than the specified tolerance, there must be omitted sources of error beyond those predicted from calibration uncertainty. The paper develops a refined bottom-up uncertainty approach for four cases of simple linear calibration: (1) inverse regression with negligible error in predictors, (2) inverse regression with non-negligible error in predictors, (3) classical regression followed by inversion with negligible error in predictors, and (4) classical regression followed by inversion with non-negligible errors in predictors. Our illustrations are of general interest, but are drawn from our experience with nuclear material assay by non-destructive assay. The main example we use is gamma spectroscopy that applies the enrichment meter principle.

Previous papers that ignore error in predictors have shown a tendency for inverse regression to have lower error variance than classical regression followed by inversion. This paper supports that tendency both with and without error in predictors. Also, the paper shows that calibration parameter estimates using error in predictor methods perform worse than without using error in predictor methods in the case of inverse regression, but perform better than without using error in predictor methods in the case of classical regression followed by inversion.

Both inverse and classical regression involve the ratio of dependent random variables; therefore, the assumed error distribution(s) will matter in parameter estimation and in uncertainty calculations. Mainly for that reason, calibration using a single predictor is distinct from simple regression, and it has not been thoroughly treated in the literature, nor in the ISO Guide to the Expression of Uncertainty in Measurements (GUM). Our refined approach is based on simulation, because we illustrate that analytical

* Corresponding author.
   E-mail address: t.burr@iaea.org (T. Burr).

approximations are not adequate when there are, for example, 10 or fewer calibration measurements, which is common in calibration applications, each consisting of measured responses from known quantities.

1. Introduction

No measurement method is complete without an estimate of its uncertainty. Uncertainty arises from errors of various types, such as random and systematic errors, and can be quantified by the variance of each known error type. In the bottom-up error variance estimation considered here, the measurement method is calibrated (as is typically required). Several guides for quantifying and expressing uncertainty in measurement are available, such as the GUM (JCGM 2008), but none gives numerical details regarding method calibration. Literature on the function fitting using least squares estimation provides analytical approximations needed in calibration applications, but little is known about their adequacy, particularly in small training sets consisting of, for example, 10 or fewer measured responses from known quantities (Tellinghuisen, 2000). Calibration is assumed to be performed using "classical forward regression followed by an inverse," in which one fits a response, such as item mass, as a function of a predictor, such as a count rate, and then inverts the fitted model to estimate the test item mass. Alternatively, we also consider that calibration can consist of "inverse regression," in which item mass is fit as a function of a predictor (Krutchoff 1967, 1969; Parker et al., 2010; Tellinghuisen, 2000).

In order to predict the variance in the errors in estimating test item masses using a bottom-up approach, calibration information, such as the number of calibration items, and the error variances in the response values can be used along with least squares regression theory; if there is non-negligible error variance in the predictors, then an approach that allows for errors in the predictors is required. We use an example in Section 7 to illustrate that available approximations for errors in predictors models are not sufficiently accurate for our goals; therefore, simulation is needed to assess calibration performance (Seber and Wild, 2003; Tellinghuisen, 2010).

If the same measurement method is used in a measurement comparison exercise and the method has been calibrated by Laboratory 1 and also by Laboratory 2, we can compare measurements of test items by both laboratories using both a top-down and bottom-up approach. In sample-exchange exercises, the top-down approach uncertainty is usually larger than predicted by the bottom-up approach. A recent paper provided a statistical partial explanation of this experience, dealing with the fact that the bottom-up approach usually requires approximate uncertainty quantification methods (Burr et al., 2014). We recognize that analytical chemists sometimes reserve the expression "top-down" for a particular empirical UQ estimation that relies on reproducibility studies involving multiple laboratories measuring the same or similar items and the objective of the experiment is to estimate the reproducibility standard deviation (ISO 21748). To clarify, we use "top-down" as a synonym for empirical, without specifying, except in particular examples, that we consider exactly what sources of the measurement error are allowed to vary. The purpose of this paper is to refine the bottom-up uncertainty estimate by using calibration information so that if there are no unknown error sources, the refined bottom-up uncertainty estimate can be brought in agreement with the top-down uncertainty estimate. Then, in practice, if the top-down uncertainty estimate is larger than the refined bottom-up uncertainty estimate, the explanation

is that there must be omitted sources of error beyond those predicted from calibration uncertainty. In non-destructive assay, item-specific deviations from calibration items and/or model is an example of such omitted error sources (item-specific deviations are beyond our scope here, but see Burr et al. (2005, 2015)). This paper develops a refined bottom-up uncertainty approach for four cases of simple linear calibration: (1) inverse regression with negligible error in predictors (EIP), (2) inverse regression with non-negligible EIP, (3) classical regression with negligible EIP, and (4) classical regression with non-negligible EIP. Our context is calibration, and to describe the four calibration options, we refer to (1) and (2) as inverse regression, and to (3) and (4) as classical regression. Previous papers that ignore EIP have shown a tendency for inverse regression to have lower error variance than classical regression followed by inversion. This paper supports that tendency both with and without EIP. Also, the paper shows that calibration parameter estimates using EIP methods perform worse than those without using EIP methods in the case of inverse regression, but perform better than the estimates without using EIP methods in the case of classical regression followed by inversion.

Both inverse and classical regression involve the ratio of dependent random variables; therefore, the assumed error distribution(s) will matter in parameter estimation and in uncertainty calculations. Mainly for that reason, calibration using a single predictor is distinct from simple regression, and it has not been thoroughly treated in the literature, nor in the GUM. Our refined approach is based on simulation, because we illustrate that analytical approximations are not adequate when there are, for example, 10 or fewer measured responses from known quantities (which is typically the case, for example, in non-destructive assay calibration applications because reference standards are expensive).

This paper is organized as follows. Section 2 gives background and describes our main example, gamma spectroscopy applying the enrichment meter principle. Section 3 describes regression assumptions, theory, and current practice in the context of calibration. Section 4 describes calibration and why standard regression theory is not fully adequate for calibration. Section 5 describes the relevant portions of the GUM. Section 6 describes measurement comparison data examples from the safeguards data evaluation section of the International Atomic Energy Agency (IAEA). Section 7 presents the four cases for which we develop a refined bottom-up uncertainty approach for simple linear calibration: (1) inverse regression with negligible error in predictors, (2) inverse regression with non-negligible error in predictors, (3) classical regression with negligible error in predictors, and (4) classical regression with non-negligible errors in predictors. Section 8 illustrates how simulation can be used to estimate variance components in a measurement error model suitable for measurement comparison data evaluation. Section 9 is a summary.

2. Background

Most assay methods rely on some type of calibration. Physics suggests that some assay methods are almost absolute (such as high-resolution gamma spectroscopy where radiation transport techniques can estimate peak efficiencies), and, in theory, would not require calibration; in practice, some of the physics